

**REPUBLIQUE ALGERIENNE DEMOCRATIQUE POPULAIRE**  
**MINISTERE D'ENSEIGNEMENT SUPERIEUR**  
**ET DE LA RECHERCHE SCIENTIFIQUE**

# **MEMOIRE**

*Présenté à l'université de BATNA*

*La faculté des sciences de l'ingénieur*  
*Département d'informatique*

**Pour l'obtention du diplôme**  
**Magister en Informatique**

**OPTION**  
**INFORMATIQUE INDUSTRIELLE**

Par  
ABDELHADI Adel

**Thème**

**Etude de l'impact Informatique de**  
**l'intégration de la langue arabe dans les**  
**téléphones mobiles**

*Soutenu le : 31/10/2004 Devant le jury composé de*

<b>M. BATOUCHE</b>	Professeur, Université de Constantine	<b>Président</b>
<b>M. BENMOHAMMED</b>	Maître de conférence, Université de Constantine	<b>Rapporteur</b>
<b>A. ZIDANI</b>	Maître de conférence, Université de Batna	<b>Examineur</b>
<b>B. BELATTAR</b>	Chargé de cours, Université de Batna.	<b>Examineur</b>

## Résumé

Ce travail présente un sujet qui entre dans le cadre du domaine d'interface homme machine. Son objectif est d'étudier l'impact informatique de l'intégration de la langue arabe dans les téléphones mobiles, afin de réaliser une interface homme machine arabe.

L'affichage correct des caractères arabes est indispensable dans une interface graphique, et puisque les caractères arabes changent leurs formes selon la position qu'ils occupent dans un mot, alors il est nécessaire de faire une analyse contextuelle sur chaque mot, pour trouver la forme correcte de chaque caractère.

La transformation de deux ou plusieurs caractères en une seule forme, demande aussi de faire un traitement particulier, comme dans le cas de la ligature **LAM-ALEF** arabe.

La langue arabe possède une direction d'écriture différente par rapport aux autres langues embarquées dans les téléphones mobiles, ce qui exige de trouver un algorithme bidirectionnel qui assure un affichage correct des messages SMS. Ces messages peuvent contenir des caractères de direction différente, de droite à gauche, de gauche à droite ou des caractères qui n'ont pas de direction. Il permet de rendre l'affichage des messages compréhensible.

### Mots clés :

Caractère Arabe, Analyse Contextuelle, Glyphe, Ligature, Algorithme Bidirectionnel, Ordre Logique, Ordre Visuel.

## Abstract

This work presents a subject which deals with man machine interfaces. Its objective is to study the data-processing impact of the integration of the Arabic language in mobile phones, in order to make an Arabic man machine interface.

The correct display of Arabic characters is essential in a graphic interface, and because Arabic characters change their forms according to the position occupied in a word, it is necessary to make a contextual analysis on each word, to find the correct form of each character.

The transformation of two or several characters in only one form, also requires to make a particular treatment, as in case of the Arabic ligature LAM-ALEF.

The Arabic language has a different direction of writing compared to other languages embedded in mobile phones, which requires to find a bi-directional algorithm that ensures a correct display of SMS messages. These messages can contain characters of different direction, right-to-left, left-to-right or characters without direction. It allows to make comprehensible the display of messages.

### Keywords:

Arabic Character, Contextual Analyzes, Glyph Shaping, Ligature, Bi-directional Algorithm, Logical Order, Visual Order.

## Sommaire

LISTE DES FIGURES.....	7
LISTE DES TABLEAUX.....	9
LISTE DES ACRONYMES .....	10
<b>introduction générale</b>	
INTRODUCTION.....	12
PLAN DU MÉMOIRE.....	13
<b>Chapitre 1 Etude Sur L'arabisation</b>	
1 INTRODUCTION.....	15
2 ARABISATION.....	15
2.1 HISTORIQUE.....	15
2.2 PRINCIPE D'ARABISATION .....	15
2.3 OBJECTIFS DE L'ARABISATION .....	16
3 INTERNATIONALISATION ET ARABISATION .....	16
3.1 INTERNATIONALISATION.....	16
3.2 ARABISATION .....	17
4 INFORMATISATION DE LA LANGUE ARABE .....	17
4.1 CODAGE INFORMATIQUE DE L'ALPHABET ARABE.....	17
4.1.1 Historique.....	17
4.1.2 De l'ASCII vers l'UNICODE .....	18
5 NOTION DE LIGATURE .....	18
6 CHIFFRES ARABES .....	19
6.1 MANIÈRE D'INSERTION DES CHIFFRES ARABES .....	19
7 DATES EN ARABE .....	20
8 CARACTÉRISTIQUES ET PROBLÈMES DE LA LANGUE ARABE.....	21
8.1 COMPOSITION DE L'ÉCRITURE ARABE .....	21
8.2 ÉCRITURE CURSIVE .....	21
8.3 DIRECTION DES CARACTÈRES ARABES.....	21
8.4 TAILLE DES CARACTÈRES ARABES.....	21
8.5 SYSTÈME DE VOCALISATION.....	21
8.5.1 VOYELLES BRÈVES (CHAACL) .....	22
8.6 FORMES DU CARACTÈRE ARABE.....	22

8.7 TYPES DE VIRGULE.....	22
8.8 FORME MINUSCULE ET MAJUSCULE .....	22
9 RÈGLES DE BASE DE L'ANALYSE CONTEXTUELLE D'UN MOT ARABE.....	23
9.1 AU DÉBUT DU MOT .....	23
9.2 AU MILIEU DU MOT .....	23
9.3 EN FIN DE MOT.....	23
10 BIDIRECTIONNALITÉ .....	23
10.1 INTRODUCTION À L'ALGORITHME BIDIRECTIONNEL .....	24
11 CONCLUSION .....	25
 <b>Chapitre 2 Unicode Et Langue Arabe</b>	
1 INTRODUCTION.....	27
2 HISTORIQUE .....	27
2.1 NORMES DE CODAGE ASCII, ANSI, ISO ET UNICODE.....	27
2.1.1 Code ASCII.....	27
2.1.2 Code ANSI.....	27
2.1.3 Codes ISO.....	28
2.1.4 Standard Unicode.....	28
3 CARACTÉRISTIQUES D'UNICODE.....	29
4 PRINCIPES DE CONCEPTION DU STANDARD UNICODE.....	30
4.1 UNIVERSALITÉ.....	30
4.2 EFFICACITÉ .....	30
4.3 CARACTÈRES ET NON GLYPHES .....	31
4.4 TEXTE BRUT.....	31
4.5 ORDRE LOGIQUE.....	32
4.6 UNIFICATION .....	32
5 ENCODAGES D'UNICODE .....	33
5.1 UTF-8.....	33
5.2 UCS-2.....	33
5.3 UTF-16.....	33
5.4 UCS-4.....	33
6 UNICODE ET LANGUE ARABE .....	34
6.1 ALGORITHME BIDI ET LANGUE ARABE.....	34
6.3 PRINCIPES DE CODAGE DES CARACTÈRES ARABES SOUS UNICODE.....	35
6.4 CARACTÈRES DE PONCTUATION SOUS UNICODE.....	35
6.5 LIAISON CURSIVE .....	35
6.5.1 Antiliant et liant.....	35
6.6 FORMES DE PRÉSENTATION ARABES A : U+FB50—U+FDFF.....	37
6.7 FORMES DE PRÉSENTATION ARABES B : U+FE70—U+FEFF .....	37

6.8 DIACRITIQUES ARABES AVEC CHASSE ET FORMES DE TATOUÏL.....	38
7 CONCLUSION .....	38

### Chapitre 3 Etude Sur Les Téléphones Mobiles

1 INTRODUCTION.....	40
2 SYSTÈME DE TÉLÉPHONE MOBILE D'UN POINT DE VUE UTILISATEUR.....	40
2.1 PARTIE MATÉRIELLE .....	40
2.2 PARTIE LOGICIELLE .....	40
3 ARABISATION DES TÉLÉPHONES MOBILES .....	41
4 DESCRIPTION DES MODULES D'UN SYSTÈME DE TÉLÉPHONE MOBILE .....	42
4.1 SAISIE DES CARACTÈRES .....	43
4.1.1 Configuration arabe (Arabic mapping) .....	43
4.1.2 Configuration syllabique.....	43
4.2 ENTRÉE PRÉDICTIVE ET ENTRÉE LIBRE .....	44
4.2.1 Entrée prédictive .....	44
4.2.2 Entrée à tape multiple(entrée libre) .....	44
4.3 API ARABE .....	44
4.3.1 Forme de glyphe (Glyph shaping).....	44
4.3.2 Algorithme bidirectionnel .....	44
4.3.3 Gestion des voyelles (CHACL).....	44
4.4 CHAMP D'ÉDITION ARABE.....	45
5 ETUDE SUR DES TÉLÉPHONES MOBILES SUPPORTANT LA LANGUE ARABE .....	46
5.1 PROMPTS .....	46
5.2 MESSAGES SMS .....	46
5.3 REMARQUES SUR LES PROMPTS ET LES MESSAGES SMS DES MODÈLES DE TÉLÉPHONE QUI SUPPORTENT L'ARABE .....	47
6 CAHIER DE CHARGE PROPOSÉ.....	48
6.1 AU NIVEAU DES PROMPTS .....	48
6.2 AU NIVEAU DES MESSAGES SMS .....	48
7 CONCLUSION .....	49

### Chapitre 4 Algorithme D'analyse Contextuelle

1 INTRODUCTION.....	51
2 DÉFINITIONS .....	51
2.1 CARACTÈRE ARABE .....	51
2.2 GLYPHE .....	51
2.3 LIGATURE.....	51
3 PRINCIPE DE L'ALGORITHME.....	52

3.1 VUE D'ENSEMBLE .....	52
3.2 TYPE DE GLYPHE ARABE .....	52
3.3 CLASSES DE LIAISON.....	52
3.4 RÈGLES GÉNÉRAL D'ALGORITHME D'ANALYSE CONTEXTUELLE.....	53
3.5 RÈGLES DE LIAISON DÉTAILLÉE DE L'ALGORITHME D'ANALYSE CONTEXTUELLE... 55	
3.5.1 <i>Au début de mot</i> .....	55
3.5.2 <i>A la fin du mot</i> .....	55
3.5.3 <i>Au milieu du mot</i> .....	56
3.6 TRAITEMENT DES LIGATURES.....	57
4 L'AUTOMATE DE L'ALGORITHME D'ANALYSE CONTEXTUELLE.....	58
5 RÈGLES DE RÉOLUTION DE LA LIGATURE.....	59
6 EXEMPLE D'IMPLÉMENTATION DE L'ALGORITHME D'ANALYSE CONTEXTUELLE.....	62
6.1 RECHERCHE D'UN MOT ARABE.....	62
6.2 APPLICATION DU GLYPH SHAPING.....	62
6.3 TRAITEMENT DE LA LIGATURE LAM / ALEF.....	62
7 ORGANIGRAMMES DE L'EXEMPLE D'IMPLÉMENTATION.....	63
<b>Chapitre 5 Algorithme D'affichage Bidirectionnel</b>	
1 INTRODUCTION.....	65
2 DÉFINITIONS .....	65
2.1 DIRECTION D'ÉCRITURE .....	65
2.2 CARACTÈRE NEUTRE.....	65
2.3 CHIFFRES ARABES .....	65
2.4 PROPRIÉTÉ MIROIR.....	65
2.5 PROPRIÉTÉ DIRECTIONNELLE.....	65
2.6 TYPE DIRECTIONNEL DE CARACTÈRE.....	66
2.7 PALIER DE CARACTÈRE.....	67
2.8 CODES DE FORMATAGE .....	67
2.9 NIVEAU D'ENCHÂSSEMENT.....	67
2.10 DIRECTION D'ENCHÂSSEMENT .....	67
2.11 ÉTAT DE FORÇAGE DIRECTIONNEL.....	68
2.12 ORDRE LOGIQUE.....	68
2.13 ORDRE VISUEL.....	68
3 ALGORITHME D'UNICODE .....	68
3.1 RÈGLES D'ÉTABLISSEMENT DES NIVEAUX D'ENCHÂSSEMENT .....	68
3.1.1 <i>Enchâssement explicites</i> .....	68
3.1.2 <i>Forçages explicites</i> .....	69
3.1.3 <i>Fin des enchâssements et des forçages</i> .....	69
3.2 RÉOLUTION DES TYPES FAIBLES .....	69

3.3	RÉSOLUTION DES TYPES NEUTRES .....	70
3.4	RÉSOLUTION DES NIVEAUX IMPLICITES.....	71
3.5	REMISE EN ORDRE DES NIVEAUX RÉSOLUS .....	72
4	L'ALGORITHME D'AFFICHAGE BIDI MOBILE .....	73
4.1	DÉFINITION DE LA DIRECTIONALITÉ D'AFFICHAGE.....	73
4.1.1	<i>Définition du type directionnel des caractères</i> .....	73
4.1.2	<i>Définition de la base de direction d'affichage</i> .....	74
4.2	DIRECTION D'AFFICHAGE DROITE À GAUCHE .....	74
4.3	DIRECTION D'AFFICHAGE GAUCHE À DROITE .....	76
5	ORGANIGRAMME DE L'ALGORITHME BIDI MOBILE .....	77
6	RÉSULTATS OBTENUS.....	78
6.1	DIRECTION D'AFFICHAGE DROITE À GAUCHE .....	78
6.2	DIRECTION D'AFFICHAGE GAUCHE À DROITE .....	79
6.3	DIRECTION D'AFFICHAGE NEUTRE.....	79
<b>Conclusion Générale</b>		
CONCLUSION .....		81
BIBLIOGRAPHIE .....		85
<b>Annexes</b>		
<b>Annexe A : Classes de liaisons</b>		
1	CLASSES DE LIAISON DES CARACTÈRES ARABES .....	90
1.1	CLASSE DES CARACTÈRES ARABES RELIÉS DES DEUX CÔTÉS. ....	90
1.2	CODES DES FORMES DE CARACTÈRES RELIÉS DES DEUX CÔTÉS .....	91
1.3	CLASSE DES CARACTÈRES ARABES RELIÉS À DROITE .....	92
1.4	CODES DES FORMES DE CARACTÈRES RELIÉS À DROITE .....	92
<b>Annexe B : Routines d'algorithme Glyph Shaping et Bidi Mobile</b>		
1	INTRODUCTION.....	93
2	ALGORITHME D'ANALYSE CONTEXTUEL .....	93
2.1	DESCRIPTION DES TYPES ET VARIABLES UTILISÉS DANS L'ALGORITHME D'ANALYSE CONTEXTUEL.....	93
2.1.1	<i>Description des constantes</i> .....	93
2.1.2	<i>Description des types complexes</i> .....	94
2.2	RECHERCHE D'UN MOT ARABE .....	95
2.3	APPLICATION DU GLYPH SHAPING .....	95
2.3.1	<i>Code de la routine DEFINITION_CLASS</i> .....	96
2.3.2	<i>Code de la routine MEDIUM_GLYPH_SHAPING</i> .....	96
2.4	TRAITEMENT DE LA LIGATURE LAM / ALEF .....	98

2.4.1 Code de la routine <i>LAM_ALEF_LIGATURE</i> .....	98
3 ALGORITHME BIDI MOBILE .....	100
3.1 BIDI_ANALYSE_STRING .....	100
3.2 DIRECTION_TYPE_CHARACT .....	101
3.3 FIND_DIRECTION .....	102
3.4 SEEK_RWORD.....	103
3.5 SEEK_LWORD.....	104
3.6 TREAT_DIRECT_RTL .....	105
3.7 TREAT_DIRECT_LTR.....	106
3.8 CODE DE LA ROUTINE BIDI_ANALYSE_STRING .....	107
3.9 CODE DE LA ROUTINE DIRECTION_TYPE_CHARACT .....	107
3.9 CODE DE LA ROUTINE DIRECTION_TYPE_CHARACT .....	108
3.10 CODE DE LA ROUTINE SEEK_RWORD.....	109
3.11 CODE DE LA ROUTINE TREAT_DIRECT_RTL.....	110
GLOSSAIRE : UNICODE .....	112